



#2

OIPF

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/10/041,016

DATE: 02/16/2002
 TIME: 13:04:10

Input Set : N:\Crf3\RULE60\10041016.txt
 Output Set: N:\CRF3\02152002\J041016.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:
 5 (i) APPLICANT: Jacobs, Kenneth
 6 McCoy, John M.
 7 Racie, Lisa A.
 8 LaVallie, Edward R.
 9 Merberg, David
 10 Treacy, Maurice
 11 Evans, Cheryl
 12 Agostino, Michael
 13 Lu, Zhijian
 14 Honjo, Tasuku
 W--> 15 Tashiro, Kei
 W--> 16 Nakamura, Tomoyuki
 18 (ii) TITLE OF INVENTION: SECRETED PROTEINS
 20 (iii) NUMBER OF SEQUENCES: 2
 22 (iv) CORRESPONDENCE ADDRESS:
 23 (A) ADDRESSEE: Genetics Institute, Inc.
 24 (B) STREET: 87 CambridgePark Drive
 25 (C) CITY: Cambridge
 26 (D) STATE: MA
 27 (E) COUNTRY: U.S.A.
 28 (F) ZIP: 02140
 30 (v) COMPUTER READABLE FORM:
 31 (A) MEDIUM TYPE: Floppy disk
 32 (B) COMPUTER: IBM PC compatible
 33 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 34 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
 36 (vi) CURRENT APPLICATION DATA:
 C--> 37 (A) APPLICATION NUMBER: US/10/041,016
 C--> 38 (B) FILING DATE: 07-Jan-2002
 44 (C) CLASSIFICATION:
 C--> 41 (vii) PRIOR APPLICATION DATA:
 42 (A) APPLICATION NUMBER: US/09/083,002
 43 (B) FILING DATE: 21-MAR-1998
 45 (viii) ATTORNEY/AGENT INFORMATION:
 46 (A) NAME: Sprunger, Suzanne A.
 47 (B) REGISTRATION NUMBER: P-41,323
 49 (ix) TELECOMMUNICATION INFORMATION:
 50 (A) TELEPHONE: (617) 498-8284
 51 (B) TELEFAX: (617) 876-5851
 54 (2) INFORMATION FOR SEQ ID NO: 1:
 56 (i) SEQUENCE CHARACTERISTICS:

ENTERED

RAW SEQUENCE LISTING

DATE: 02/16/2002

PATENT APPLICATION: US/10/041,016

TIME: 13:04:10

Input Set : N:\Crf3\RULE60\10041016.txt

Output Set: N:\CRF3\02152002\J041016.raw

```

57      (A) LENGTH: 2362 base pairs
58      (B) TYPE: nucleic acid
59      (C) STRANDEDNESS: double
60      (D) TOPOLOGY: linear
62      (ii) MOLECULE TYPE: cDNA
63      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
64      TAGCTTGGCA CGAGGGGACC CCGGCGTCT CCCGCTGTCC TCTCCACGAC TCGCTCGGCC      60
65      CCTCTGGAAT AAAACACCCG CGAGCCCCGA GGGCCCAGAG GAGGCCGACG TGCCCCGAGCT      120
66      CCTCCGGGGG TCCCGCCCGC GAGCTTTCTT CTCGCCCTCG CATCTCCTCC TCGCGCGTCT      180
67      TGGACATGCC AGGAATAAAA AGGATACTCA CTGTTACCAT TCTGGCTCTC TGTCTTCCAA      240
68      GCCCTGGGAA TGCACAGGCA CAGTGCACGA ATGGCTTTGA CCTGGATCGC CAGTCAGGAC      300
69      AGTGTTTAGA TATTGATGAA TGCCGAACCA TCCCGGAGGC CTGCCGAGGA GACATGATGT      360
70      GTGTTAACCA AAATGGCGGG TATTTATGCA TTCCCCGGAC AAACCCCTGTG TATCGAGGGG      420
71      CCTACTCGAA CCCCTACTCG ACCCCCTACT CAGGTCCGTA CCCAGCAGCT GCCCCACCAC      480
72      TCTCAGCTCC AAACATATCC ACAGATCTCCA GGCTCTTAT ATGCCGCTTT GGATACCAGA      540
73      TGGATGAAAG CAACCAATGT GTGGATGTGG ACGAGTGTGC AACAGATTCC CACCACTGAT      600
74      ACCCCACCCA GATCTGCATC AATACTGAAG GCGGGTACAC CTGCTCCTGC ACCGACGGA      660
75      ATTGGCTTCT GGAAGGCCAG TGCTTAGACA TTGATGAATG TCGCTATGGT TACTGCCAGC      720
76      AGCTCTGTGC GAATGTTCTT GGATCCTATT CTTGTACATG CAACCTCGGT TTTACCCCTA      780
77      ATGAGGATGG AAGGTCTTGC CAAGATGTGA ACGAGTGTGC CACCGAGAAC CCCTGCGTGC      840
78      AAACCTGCGT CAACACCTAC GGCTCTTTCA TCTGCCGCTG TGACCCAGGA TATGAACCTG      900
79      AGGAAGATGG CGTTTCATTG AGTGATATGG ACGAGTGCAG CTTCTCTGAG TTCTCTGCC      960
80      AACATGAGTG TGTGAACCAAG CCCGGCACAT ACTTCTGCTC CTGCCCTCCA GGCTACATCC      1020
81      TGCTGGATGA CAACCGAAGC TGCCAAGACA TCAACGAATG TGAGCACAGG AACCAACGCT      1080
82      GCAACCTGCA GCAGACGTGC TACAATTTAC AAGGGGGCTT CAAATGCATC GACCCCATCC      1140
83      GCTGTGAGGA GCCTTATCTG AGGATCAGTG ATAACCGCTG TATGTGTCTT GCTGAGAAC      1200
84      CTGGCTGCAG AGACCAGCCC TTTACCATCT TGTACCGGGA CATGGACGTG GTGTCAAGAC      1260
85      GCTCCGTTCG CGCTGACATC TTCCAAATGC AAGCCACGAC CCGCTACCCT GGGGCCCTATT      1320
86      ACATTTTCCA GATCAAATCT GGAATGAGG GCAGAGAATT TTACATGCGG CAAACGGGCC      1380
87      CCATCAGTGC CACCTGGTG ATGACACGCC CCATCAAAGG GCCCCGGGAA ATCCAGCTGG      1440
88      ACTTGGAATG GATCACTGTC AACACTGTCA TCAACTTCAG AGGCAGCTCC GTGATCCGAC      1500
89      TGCGGATATA TGTGTCGACG TACCCATTCT GAGCCTCGGG CTGGAGCCTC CGACGCTGCC      1560
90      TCTCATTGGC ACCAAGGGAC AGGAGAAGAG AGGAAATAAC AGAGAGAATG AGAGCGACAC      1620
91      AGACGTTAGG CATTTCTGTC TGAACGTTTC CCCGAAGAGT CAGCCCCGAC TTCCTGACTC      1680
92      TCACCTGTAC TATTGCAGAC CTGTCAACCT GCAGGACTTG CCACCCCGAG TCCCTATGAT      1740
93      ACAGTTATCA AAAAGTATTA TCATTGCTCC CCTGATAGAA GATTGTTGGT GAATTTTCAA      1800
94      GGCCTTCAGT TTAATTTCCAC TATTTTCAAA GAAATAGAT TAGGTTTTCG GGGTCTGAG      1860
95      TCTATGTTCA AAGACTGTGA ACAGCTTGCT GTCATTCTCT CACTCTTCC ACTCCTTCTC      1920
96      TCACTGTGTT ACTGCTTTGC AAAGACCCGG GAGCTGGCGG GGAACCTTGG GAGTAGCTAG      1980
97      TTTGCTTTTT GCGTACACAG AGAAGGCTAT GTAAACAAAC CACAGCAGGA TCGAAGGGTT      2040
98      TTTAGAGAAT GTGTTTCAAA ACCATGCGCTG GTATTTTCAA CCATAAAAGA AGTTTCAGTT      2100
99      GTCCCTTAAAT TTGTATAACG GTTTAATTCT GTCTGTGTCA TTTTGAGTAT TTTTAAAAAA      2160
100      TATGTCGTAG AATTCCTTCG AAAGGCCCTC AGACACATGC TATGTTCTGT CTTCCCAAAC      2220
101      CCAGTCTCCT CTCCATTTTA GCCCAGTGTG TTCTTTGAGG ACCCCTTAAT CTTGCTTTCT      2280
102      TTAGAATTTT TACCCAATTG TATGGAATG CAGAGGTCTC CAACTGATT AAATATTTGA      2340
103      ATGAAAAAAA AAAAAAAA AA
104      (2) INFORMATION FOR SEQ ID NO: 2:
105      (i) SEQUENCE CHARACTERISTICS:
106      (A) LENGTH: 448 amino acids

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/041,016

DATE: 02/16/2002

TIME: 13:04:10

Input Set : N:\Crf3\RULE60\10041016.txt

Output Set : N:\CRF3\02152002\J041016.raw

```

153         (B) TYPE: amino acid
154         (C) STRANDEDNESS: Not Relevant
155         (D) TOPOLOGY: linear
156
157 (ii) MOLECULE TYPE: protein
158
159 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
160 Met Pro Gly Ile Lys Arg Ile Leu Thr Val Thr Ile Leu Ala Leu Cys
161      1      5      10      15
162 Leu Pro Ser Pro Gly Asn Ala Gln Ala Gln Cys Thr Asn Gly Phe Asp
163      20      25      30
164 Leu Asp Arg Gln Ser Gly Gln Cys Leu Asp Ile Asp Glu Cys Arg Thr
165      35      40      45
166 Ile Pro Glu Ala Cys Arg Gly Asp Met Met Cys Val Asn Gln Asn Gly
167      50      55      60
168 Gly Tyr Leu Cys Ile Pro Arg Thr Asn Pro Val Tyr Arg Gly Pro Tyr
169      65      70      75      80
170 Ser Asn Pro Tyr Ser Thr Pro Tyr Ser Gly Pro Tyr Pro Ala Ala Ala
171      85      90      95
172 Pro Pro Leu Ser Ala Pro Asn Tyr Pro Thr Ile Ser Arg Pro Leu Ile
173      100     105     110
174 Cys Arg Phe Gly Tyr Gln Met Asp Glu Ser Asn Gln Cys Val Asp Val
175      115     120     125
176 Asp Glu Cys Ala Thr Asp Ser His Gln Cys Asn Pro Thr Gln Ile Cys
177      130     135     140
178 Ile Asn Thr Glu Gly Gly Tyr Thr Cys Ser Cys Thr Asp Gly Tyr Trp
179      145     150     155     160
180 Leu Leu Glu Gly Gln Cys Leu Asp Ile Asp Glu Cys Arg Tyr Gly Tyr
181      165     170     175
182 Cys Gln Gln Leu Cys Ala Asn Val Pro Gly Ser Tyr Ser Cys Thr Cys
183      180     185     190
184 Asn Pro Gly Phe Thr Leu Asn Glu Asp Gly Arg Ser Cys Gln Asp Val
185      195     200     205
186 Asn Glu Cys Ala Thr Glu Asn Pro Cys Val Gln Thr Cys Val Asn Thr
187      210     215     220
188 Tyr Gly Ser Phe Ile Cys Arg Cys Asp Pro Gly Tyr Glu Leu Glu Glu
189      225     230     235     240
190 Asp Gly Val His Cys Ser Asp Met Asp Glu Cys Ser Phe Ser Glu Phe
191      245     250     255
192 Leu Cys Gln His Glu Cys Val Asn Gln Pro Gly Thr Tyr Phe Cys Ser
193      260     265     270
194 Cys Pro Pro Gly Tyr Ile Leu Leu Asp Asp Asn Arg Ser Cys Gln Asp
195      275     280     285
196 Ile Asn Glu Cys Glu His Arg Asn His Thr Cys Asn Leu Gln Gln Thr
197      290     295     300
198 Cys Tyr Asn Leu Gln Gly Gly Phe Lys Cys Ile Asp Pro Ile Arg Cys
199      305     310     315     320
200 Glu Glu Pro Tyr Leu Arg Ile Ser Asp Asn Arg Cys Met Cys Pro Ala
201      325     330     335
202 Glu Asn Pro Gly Cys Arg Asp Gln Pro Phe Thr Ile Leu Tyr Arg Asp
203      340     345     350

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/041,016

DATE: 02/16/2002

TIME: 13:04:10

Input Set : N:\Crf3\RULE60\10041016.txt

Output Set: N:\CRF3\02152002\J041016.raw

230	Met	Asp	Val	Val	Ser	Gly	Arg	Ser	Val	Pro	Ala	Asp	Ile	Phe	Gln	Met
231			355						360				365			
233	Gln	Ala	Thr	Thr	Arg	Tyr	Pro	Gly	Ala	Tyr	Tyr	Ile	Phe	Gln	Ile	Lys
234			370				375					380				
236	Ser	Gly	Asn	Glu	Gly	Arg	Glu	Phe	Tyr	Met	Arg	Gln	Thr	Gly	Pro	Ile
237			385			390					395				400	
239	Ser	Ala	Thr	Leu	Val	Met	Thr	Arg	Pro	Ile	Lys	Gly	Pro	Arg	Glu	Ile
240				405						410					415	
242	Gln	Leu	Asp	Leu	Glu	Met	Ile	Thr	Val	Asn	Thr	Val	Ile	Asn	Phe	Arg
243				420					425					430		
245	Gly	Ser	Ser	Val	Ile	Arg	Leu	Arg	Ile	Tyr	Val	Ser	Gln	Tyr	Pro	Phe
246			435					440					445			

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/041,016

DATE: 02/16/2002

TIME: 13:04:11

Input Set : N:\Crf3\RULE60\10041016.txt

Output Set: N:\CRF3\02152002\J041016.raw

L:37 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:38 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:41 M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]
L:15 M:259 W: Allowed number of lines exceeded, (i) APPLICANT:
L:16 M:259 W: Allowed number of lines exceeded, (i) APPLICANT: